Exponential Separations for One-Way Quantum Communication Complexity, with Applications to Cryptography

Dmitry Gavinsky*
IQC, University of Waterloo

Julia Kempe†
CNRS & LRI, Univ. Paris-Sud, Orsay
and School of CS, Tel Aviv Univ.

Iordanis Kerenidis†
CNRS & LRI, Univ. Paris-Sud, Orsay

Ran Raz‡
Faculty of Maths, Weizmann

Ronald de Wolf§
CWI, Amsterdam

ABSTRACT

We give an exponential separation between one-way quantum and classical communication protocols for two partial Boolean functions, both of which are variants of the Boolean Hidden Matching Problem of Bar-Yossef et al. Earlier such an exponential separation was known only for a relational version of the Hidden Matching Problem. Our proofs use the Fourier coefficients inequality of Kahn, Kalai, and Linial. We give a number of applications of this separation. In particular, in the bounded-storage model of cryptography we exhibit a scheme that is secure against adversaries with a certain amount of classical storage, but insecure against adversaries with a similar (or even much smaller) amount of *quantum* storage; in the setting of privacy amplification, we show that there are strong extractors that yield a classically secure key, but are insecure against a quantum adversary.

Categories and Subject Descriptors

E.4 [**Coding and information theory**]: Formal models of communication; F.1.3 [**Computation by Abstract Devices**]: Complexity Measures and Classes—*Relations among complexity measures*

General Terms

Algorithms, Theory

*Supported in part by Canada's NSERC.
†Supported by ACI Sécurité Informatique SI/03 511 and ANR AlgoQP grants, and also by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848.
‡Part of this work was done when this author visited Microsoft Research, supported by ISF and BSF grants.
§Supported by a Veni grant from the Netherlands Organization for Scientific Research (NWO) and also partially supported by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848.

STOC'07, June 11–13, 2007, San Diego, California, USA.

Keywords

quantum, communication complexity, cryptography

1. INTRODUCTION

One of the main goals of quantum computing is to find problems where quantum computers are much faster (or otherwise better) than classical computers. Preferably exponentially better. The most famous example, Shor's quantum factoring algorithm [31], is a separation only if one is willing to believe that efficient factoring is impossible on a classical computer—proving this would, of course, imply P≠NP. One of the few areas where one can establish *unconditional* exponential separations is communication complexity.

Communication complexity is a central model of computation, first defined by Yao [36], that has found applications in many areas [18]. Two parties, Alice with input x and Bob with input y, collaborate to solve a computational problem that depends on both x and y. Their goal is to do this with minimal communication. The problem to be solved could be a function $f(x, y)$ or some relational problem where for each x and y, several outputs are valid. The protocols could be *interactive (two-way)*, in which case Alice and Bob take turns sending mesages to each other; *one-way*, in which case Alice sends a single message to Bob who then determines the output; or *simultaneous*, where Alice and Bob each pass one message to a third party (the *referee*) who determines the output. The (bounded-error) *communication complexity* of the problem is the worst-case communication of the best protocol that gives (for every input x and y) a correct output with probability at least $1-\varepsilon$, for a fixed $\varepsilon \in [0, 1/2)$.

Allowing the players to use *quantum* instead of classical resources can reduce the communication complexity significantly. Examples of problems where quantum communication gives exponential savings were given by Buhrman, Cleve, and Wigderson for one-way and interactive protocols with zero error probability [5]; by Raz for bounded-error interactive protocols [28]; and by Buhrman, Cleve, Watrous, and de Wolf for bounded-error simultaneous protocols [6]. The first two problems are partial Boolean functions, while the third one is a total Boolean function (however, that separation doesn't hold in the presence of public coins). In fact, whether there exists a superpolynomial separation for a total Boolean function in the presence of public coins is one of the main open questions in the area.

Moreover, Bar-Yossef, Jayram, and Kerenidis [8] showed an exponential separation for one-way protocols and simultaneous protocols with public coins, but they only achieve

this for a relational problem, called the *Hidden Matching Problem* (HMP). This problem can be solved efficiently by one quantum message of $\log n$ qubits, but classical one-way protocols need to send nearly $\sqrt{n}$ bits to solve it. However, Boolean functions are much more natural objects than relations both in the model of communication complexity and in the cryptographic settings that we consider later in this paper. Bar-Yossef et al. stated a Boolean version of their problem (partial function) and conjectured that the same quantum-classical gap holds for this problem as well.

1.1 Exponential separation for NPM

We prove tight bounds for the bounded-error one-way communication complexity of a slight variant of Boolean Hidden Matching, which we call the *Noisy Perfect Matching problem* (NPM). Precise definitions are in Section 2.

THEOREM 1. *The classical bounded-error one-way communication complexity of the Noisy Perfect Matching problem is* $R^1_\varepsilon(\mathrm{NPM}) = \Theta(\sqrt{n})$*, while the quantum bounded-error one-way complexity is* $Q^1_\varepsilon(\mathrm{NPM}) = \Theta(\log n)$.

This is the first exponential separation between quantum and classical one-way communication complexity for a partial Boolean function. Our $\Omega(\sqrt{n})$ lower bound is proved using the Fourier coefficients inequality of Kahn, Kalai, and Linial [16], which is a special case of the Bonami-Beckner inequality [9, 7]. Fourier analysis was previously used in communication complexity by Raz [27] and Klauck [17].

1.1.1 Application: streaming model

In the *streaming model* of computation, the input is given as a stream of bits and the algorithm is supposed to compute or approximate some function of the input, having only space of size S available. See for instance [3, 24]. There is a well-established connection between one-way communication complexity and the streaming model: if we view the input as consisting of two parts x and y, then the content of the memory after x has been processed, together with y, contains enough information to compute $f(x, y)$. Hence, a space-S streaming algorithm for f implies a one-way protocol for f of communication S. The classical lower bound for our communication problem, together with the observation that our quantum protocol can be implemented in the streaming model, implies a separation between the quantum and classical streaming model: there is a partial Boolean function f that can be computed in the streaming model with small error probability using quantum space of $O(\log n)$ qubits, but requires $\Omega(\sqrt{n})$ bits if the space is classical.

Le Gall [12] constructed a problem that can be solved in the streaming model using $O(\log n)$ qubits of space, while any classical algorithm needs $\Omega(n^{1/3})$ classical bits. His $\log n$-vs-$n^{1/3}$ separation is a bit smaller than our $\log n$-vs-$\sqrt{n}$, but his separation is for a *total* Boolean function while ours is only partial (i.e., requires some promise on the input). Le Gall's result predates ours, though we only learned about it after finishing our paper. While Le Gall's separation holds only in the streaming model variant where the bits arrive in order, ours holds in the more general model, where we allow the input bits to arrive in any order.

1.2 A variant with links to cryptography

Our next result deals with another variant of the Boolean Hidden Matching Problem, called the α-Partial Matching problem (αPM), which is parametrized by some value α as defined in Section 2. The ability to vary this parameter α will be important for some of our applications. For this variant we can also establish an exponential gap:

THEOREM 2. *For* $\alpha \in [0, O(1/\sqrt{\log n})]$*, the classical bounded-error one-way communication complexity of* α*-Partial Matching is* $R^1_\varepsilon(\alpha\mathrm{PM}) = \Theta(\sqrt{n/\alpha})$*; the quantum bounded-error one-way complexity is* $Q^1_\varepsilon(\alpha\mathrm{PM}) = O(\log(n)/\alpha)$.

For instance for $\alpha \approx 1/\sqrt{\log n}$ the separation is $(\log n)^{3/2}$ qubits versus $\sqrt{n}(\log n)^{1/4}$ classical bits. The quantum protocol for αPM is less efficient than the quantum protocol for NPM ($(\log n)^{3/2}$ vs $\log n$ qubits), but the latter has bounded error while the former can be made to run with error probability 0 with *expected* communication $O(\log(n)/\alpha)$.

For the cryptographic applications below, it is crucial that the proof of this second separation actually shows that if Alice's message was too short, then Bob has hardly any information about a certain string z that can be computed from x given also Bob's input y. That is, from his perspective (given y and Alice's message) this string z is almost uniformly distributed. Our proof uses a result of Talagrand [32] (which is easy to derive from, again, the KKL inequality, though Talagrand himself proves it differently) and a large deviation inequality for martingales [23].

1.2.1 Application: the bounded storage model

Our second proof is closely related to the *bounded storage model* in cryptography. It was introduced by Maurer [22] with the aim of implementing information-theoretically secure *key expansion*. In this setting, a large random variable X is publicly but only temporarily available. Alice and Bob use a shared secret key Y to extract an additional key $Z(X, Y)$ from X. The secret key Y remains hidden from the adversary during this extraction phase, but may be revealed later. The adversary is assumed to have only a bounded amount of storage and as a result his information about Z is limited *even* if he learns the secret key Y afterwards. "Limited information" means that the distribution on $Z(X, Y)$ is η-close to uniform even when conditioned on Y and on the information about X that the adversary stored in his memory, for small security parameter $\eta \in [0, 1]$ (the smaller the better). Aumann, Ding, and Rabin [2] were the first to prove a bounded-storage scheme secure, and essentially tight constructions have subsequently been found [11, 21, 33]. It is an important open question whether any of these constructions remains secure if the adversary can store *quantum information*. One may even conjecture that a bounded-storage protocol secure against classical adversaries with a certain amount of memory, should be roughly as secure against *quantum* adversaries with roughly the same memory bound. After all, Holevo's theorem [14] informally says that k qubits cannot contain more information than k classical bits. Using the stronger statement on the uniformity of Z shown in our second separation we refute the latter conjecture.

The link to one-way communication comes from viewing Alice's input as the temporarily available randomness X, while Bob's input takes the role of the secret key Y. Alice's message $m(X)$ (which she sends *without* knowing Y) represents the stored information of the adversary about the string X before he learns the key Y. Our lower bound proof for one-way communication shows that Bob cannot learn

much about a certain αn-bit string $Z(X,Y)$ if Alice's message is too short. This can be translated back to show that an adversary cannot learn much about the extracted key Z if his storage is too small. Our result gives the first example of a bounded-storage protocol where the extracted key can be made η-secure[1] against a classical adversary (for any constant η) but becomes completely insecure against a quantum adversary of the same or even much smaller memory size.

THEOREM 3. *Let $\eta \in [0,1]$ and $\alpha \in [0, O(\sqrt{\eta/\log n})]$. The extracted αn-bit string in the bounded-storage protocol derived from the αPM problem is η-secure against a classical adversary with memory bound $O(\sqrt{\eta^3 n/\alpha})$, while for every positive integer $k \ll \alpha n$ it is at most $(1-2^{-k})$-secure against an adversary with $O(k\log(n)/\alpha)$ qubits.*

Note that normally in cryptography one wants η-security for exponentially small η. Our classical bounded-storage scheme is not secure in that strong sense, but it is secure for any constant η of our choice. In fact, by choosing α appropriately, we can make η inverse-polynomially small.

It should be noted that the bounded-storage protocol derived from αPM—though provably secure against classical adversaries—is not terribly useful. Usually one wants the initial key Y to be *much* smaller than the extracted key Z, and this is actually achieved by the classical schemes cited above. In our scheme the initial key Y is actually *longer* than the final key Z. It can still be used for key expansion, where one expands a secure key Y to a longer secure key $Y, Z(X,Y)$. Though it would be interesting to find a constructive example with much shorter initial key, the main point of our result here is to givee an example of a classically-secure scheme that is insecure against quantum.

1.2.2 Application: extractors, privacy amplification

The proof of our second separation is also closely related to the notion of strong *randomness extractors*. There the task is to extract almost uniform randomness from an *imperfect* (i.e. non-uniform) source of randomness X with the help of an independent uniform seed Y. In other words, the output of an extractor is a random variable $Z(X,Y)$, such that the pair $(Y, Z(X,Y))$ is close to uniform. The main parameters of an extractor are the length of the uniformly random string Y, and the randomness of the imperfect source, which is measured by the *min-entropy* of the source.

Extractors have found numerous applications in computer science, in particular in complexity theory and cryptography. One important application is that of *privacy amplification*, which was introduced in [4, 15]. In this setting, Alice and Bob start with a shared random variable X about which the adversary has some partial information $m(X)$ and their goal is to generate a secret key Z about which the adversary has very little information. They can achieve this by communicating an independent uniform seed Y over an insecure channel, and using a strong extractor to generate the key $Z(X,Y)$. Assuming a certain upper bound on the number of bits of $m(X)$, the key $Z(X,Y)$ is secure even if the adversary has full knowledge of Y.

[1]This means that the distribution on Z is η-close to uniform, conditioned on Y. Formally, $\mathrm{E}_y[d(Z(X,Y), U \mid Y=y)] \le \eta$, where $d(p,q) = \frac{1}{2}\sum_x |p(x)-q(x)|$ denotes total variation distance, U is the uniform distribution, and expectation is taken uniformly over all possible values y of Y.

Extractors and privacy amplification can also be considered in the *quantum* case where the prior partial information about the string X is a quantum state. Our communication result implies that there exist extractors which yield a classically secure key, but that are insecure against a quantum adversary. More specifically, one can think of Alice's input X as the shared random variable, her message $m(X)$ as the prior partial information of the adversary about X, and Bob's input Y as the independent uniform seed. Our lower bound shows that in the classical setting, the αn-bit string $Z(X,Y)$ is close to uniform even if the size of the classical prior information $m(X)$ is as large as $O(\sqrt{n})$. However, in the quantum setting the key becomes insecure even if the quantum prior information is of size only $poly(\log n)$.

The dependence of the security on whether the adversary has quantum or classical memory is quite surprising, particularly in light of the following two facts: first, privacy amplification based on two-universal hashing provides exactly the same security against classical and quantum adversaries. The length of the key that can be extracted is given by the min-entropy both in the classical ([4, 15]) and the quantum case ([30], [29, Ch. 5]). Second, König and Terhal [19] have recently shown that for protocols that extract just *one* bit, the level of security against a classical and a quantum adversary (with the same information bound) is again comparable.

1.2.3 Application: simulations of quantum protocols

Another application of our second separation is in the context of simulating one-way quantum communication protocols by one-way classical protocols. As noted by Aaronson [1, Section 5], our Theorem 2 implies that his general simulation of bounded-error one-way quantum protocols by deterministic one-way protocols

$$D^1(f) = O(mQ^1_\varepsilon(f)\log Q^1_\varepsilon(f)),$$

is tight up to a polylog factor. Here m is the length of Bob's input. This simulation works for any partial Boolean function f. Taking f to be our αPM for $\alpha \approx 1/\sqrt{\log n}$, one can show that $D^1(f) \approx n$, $m \approx \alpha n \log n \approx n\sqrt{\log n}$, $Q^1_\varepsilon(f) \approx (\log n)^{3/2}$. It also implies that his simulation of quantum bounded-error one-way protocols by classical ones

$$R^1_\varepsilon(f) = O(mQ^1_\varepsilon(f)),$$

cannot be much improved. In particular, the product on the right cannot be replaced by the sum: if we take $f = \alpha$PM with $\alpha = 1/\sqrt{n}$, then by Theorem 2 we have $R^1_\varepsilon(f) \approx n^{3/4}$, $m \approx \sqrt{n}\log n$, and $Q^1_\varepsilon(f) \approx \sqrt{n}\log n$.

Remark. Our results can be modified to give a separation in the simultaneous message passing model between classical communication with shared entanglement and classical communication with shared randomness. Earlier, such a separation was known only for a relational problem [13].

2. THE PROBLEMS AND UPPER BOUNDS

We assume basic knowledge of quantum computation [25] and (quantum) communication complexity [18, 34].

Before giving the definitions of our two variants of the Boolean Hidden Matching Problem, we fix some notation. Part of Bob's input will be a sequence M of αn disjoint edges $e_1 = (i_1, j_1), \ldots, e_{\alpha n} = (i_{\alpha n}, j_{\alpha n})$ from $[2n]$, which

we call an α-matching. If $\alpha < 1$, the matching is *partial*, if $\alpha = 1$ the matching is *perfect*. We can also view an α-matching on $[2n]$ as an $(\alpha n \times 2n)$ matrix M over $GF(2)$, where each column corresponds to a number in $[2n]$ and the ℓ-th row corresponds to the ℓ-th edge of the matching. In other words, if the ℓ-th edge of the matching is (i_ℓ, j_ℓ), then the ℓ-th row of the matrix contains two 1's at the positions i_ℓ and j_ℓ and 0's elsewhere. Let $x \in \{0,1\}^{2n}$. Then the product Mx is an αn-bit string $z = z_1, \ldots, z_\ell, \ldots z_{\alpha n}$ where $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$. Denote by $h(\cdot,\cdot)$ the Hamming distance function and by $h(\cdot)$ the Hamming weight function.

Using this notation, we introduce the two partial functions we study, which differ only in the parameter α and in the promise. We call them the *Noisy Perfect Matching (*NPM*)* and the *α-Partial Matching (α*PM*)* respectively.

Alice: $x \in \{0,1\}^{2n}$
Bob: an α-matching M on $[2n]$ and a string $w \in \{0,1\}^{\alpha n}$ ($\alpha = 1$ for NPM)
a) Promise for NPM: $\exists\, b$ such that $h(Mx \oplus b^n, w) \leq n/3$
b) Promise for αPM: $\exists\, b$ such that $w = Mx \oplus b^{\alpha n}$
Function value: b

We can draw an analogy with two kinds of noise in transmission channels. In the Noisy Perfect Matching problem, Bob's input w results from the string Mx or $Mx \oplus 1^n$ after at most a 1/3-fraction of the bits have been "corrupted". In the α-Partial Matching problem, Bob's input w can be viewed as an n-bit string resulting from a perfect matching followed by the "erasure" of a $(1-\alpha)$-fraction of the bits. For the communication complexity separation by the α-Partial Matching problem, we could fix α to an appropriate value, however, the general result is useful for our applications.

Quantum upper bounds. There is an easy $O(\log(n)/\alpha)$ protocol for both problems. Alice sends a uniform superposition of $x = x_1 \ldots x_{2n}$: $|\psi\rangle = \frac{1}{\sqrt{2n}} \sum_{i=1}^{2n} (-1)^{x_i} |i\rangle$. Bob completes his αn edges to a perfect matching in an arbitrary way, and measures with the corresponding set of n 2-dimensional projectors. With probability α he will get one of the edges $e_\ell = (i_\ell, j_\ell)$ of his input. The state then collapses to $(-1)^{x_{i_\ell}} |i_\ell\rangle + (-1)^{x_{j_\ell}} |j_\ell\rangle$, from which Bob can obtain $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$ by an appropriate measurement.

In NPM, Bob outputs $z_\ell \oplus w_\ell$. The protocol is correct with probability at least 2/3, and by repeating $O(\log(1/\varepsilon))$ times we can achieve correctness $1-\varepsilon$ for any constant $\varepsilon > 0$.

In the case of αPM, Bob can obtain the bit $b = z_\ell \oplus w_\ell$ with certainty if he has measured one of his edges (which happens with probability α), otherwise he claims ignorance. Note that this protocol has so-called "zero-sided error" (Bob knows when he didn't learn the bit b) and the success can be boosted to $1 - \varepsilon$ given $O(\log(1/\varepsilon)/\alpha)$ copies of the state.

The above protocol for αPM can be repeated k times in parallel: if Bob is given $O(k/\alpha)$ copies of $|\psi\rangle$, then with high probability (at least while $k \ll \alpha n$) he can learn k bits of z.

Classical upper bounds. We sketch an $O(\sqrt{n/\alpha})$ *classical* upper bound for both functions. Suppose Alice uniformly picks a subset of $d \approx \sqrt{n/\alpha}$ bits of x to send to Bob. By the birthday paradox, with high probability Bob will have both endpoints of at least one of his αn edges and so he can compute the function value b with good probability. In this protocol Alice would need to send about $d \log n$ bits to Bob, since she needs to describe the d indices as well as their bitvalues. However, by Newman's Theorem [26], Alice can actually restrict her random choice to picking one out of $O(n)$ possible d-bit subsets, instead of one out of all $\binom{2n}{d}$ possible subsets. Hence $d + O(\log n)$ bits suffice.

In Section 3.1 we show that for NPM the classical upper bound of $O(\sqrt{n})$ is optimal, and in Section 4 we show for αPM that for $\alpha \ll 1/\sqrt{\log n}$ the classical upper bound of $O(\sqrt{n/\alpha})$ is optimal. Choosing $\alpha \approx 1/\sqrt{\log n}$ gives a function that can be computed with $O((\log n)^{3/2})$ qubits of one-way communication, but needs at least $\Omega(\sqrt{n}(\log n)^{1/4})$ classical bits of communication, which gives the exponential quantum-classical separation for αPM.

3. LOWER BOUND FOR NPM

We prove a lower bound on classical communication with shared randomness for the problems of the previous section in two different ways. Let us first describe what is common among both proofs. By the Yao principle [35], it suffices to prove a lower bound for *deterministic* protocols under some "hard" input distribution. For both problems we choose a distribution that is uniform on the x's, the matchings M, and b. In the case of αPM this fixes Bob's second input $w = Mx \oplus b^{\alpha n}$. For the Noisy Perfect Matching problem we will in addition fix a distribution on the n-bit string w in the following way: independently choose each bit w_ℓ such that $\Pr[w_\ell = (Mx)_\ell \oplus b] = 3/4$. In other words, we can think of w as a *noisy* version of $Mx \oplus b^n = z \oplus b^n$ where each bit of $z \oplus b^n$ is flipped with probability 1/4. Note that if (x, M, b, w) are picked according to this distribution, then the probability that the Hamming distance $h(Mx \oplus b^n, w)$ is more than $n/3$, is exponentially small. Hence, any probabilistic protocol for NPM with error ε' gives a deterministic protocol for this distribution with distributional error $\varepsilon' + o(1)$. Therefore, for the rest of the proof we use this distribution.

Suppose we have a classical deterministic one-way protocol with c bits and error probability at most ε under this distribution for either NPM or αPM. This protocol partitions the set of 2^{2n} x's into 2^c sets $A_1, \ldots, A_{2^c}$, one for each possible message. Note that on average, these sets have size 2^{2n-c}. Moreover, at most an η-fraction of all $x \in \{0,1\}^{2n}$ can sit in sets of size $\leq 2^{2n-c-\log(1/\eta)}$. In particular, at least half of the x's must occur in sets of size at least 2^{2n-c-1}. Hence there must be at least one set A that contains at least 2^{2n-c-1} x's and has error at most 2ε, otherwise the overall error would be larger than ε. Hereafter, we analyze this A.

3.1 Fourier analysis of NPM

Our proof for NPM directly bounds Bob's probability to learn b. In order to learn b, Bob needs to determine whether his string w comes from a noisy version of $Mx \oplus 0^n$ or of $Mx \oplus 1^n$. We upper bound the total variation distance between these two distributions using Fourier analysis. This gives an upper bound on the size of A, and hence a lower bound on the communication c. We begin by providing a few standard definitions from Fourier analysis.

For functions $f, g : \{0,1\}^n \to \mathbb{R}$ we define their inner product and the ℓ_1, ℓ_2 norms by $\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) g(x)$, $||f||_1 = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} |f(x)|$, $||f||_2^2 = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} |f(x)|^2$. Note that $||f||_2 \geq ||f||_1$ by Cauchy-Schwarz. The Fourier transform of f is a function $\hat{f} : \{0,1\}^n \to \mathbb{R}$ with $\hat{f}(s) = \langle f, \chi_s \rangle = \frac{1}{2^n} \sum_{y \in \{0,1\}^n} f(y) \chi_s(y)$, where $\chi_s : \{0,1\}^n \to \mathbb{R}$ is the character $\chi_s(y) = (-1)^{y \cdot s}$ with "$\cdot$" being the scalar product; $\hat{f}(s)$ is the Fourier coefficient of f corresponding

to s. We have the following relation between f and $\hat{f}$: $f = \sum_{s\in\{0,1\}^n} \hat{f}(s)\chi_s$. The convolution $f * g : \{0,1\}^n \to \mathbb{R}$ for $f, g : \{0,1\}^n \to \mathbb{R}$ is $f * g(w) = \frac{1}{2^n}\sum_{y\in\{0,1\}^n} f(y \oplus w)g(y)$. Note that with this definition we have $\widehat{(f * g)}(s) = \hat{f}(s)\cdot\hat{g}(s)$. We also use Parseval's identity and the KKL lemma.

LEMMA 4 (PARSEVAL'S IDENTITY). *For every function $f : \{0,1\}^n \to \mathbb{R}$, $||f||_2^2 = \sum_{s\in\{0,1\}^n}(\hat{f}(s))^2$.*

LEMMA 5 ([16]). *Let f be a function $f : \{0,1\}^n \to \{-1,0,1\}$. Let $t = |\{x \mid f(x) \neq 0\}|/2^n$ be the uniform probability that $f \neq 0$. Then for every $\delta \in [0,1]$ we have*

$$\sum_{s\in\{0,1\}^n} \delta^{h(s)}(\hat{f}(s))^2 \leq t^{\frac{2}{1+\delta}}.$$

PROOF OF THEOREM 1. Following the lead of Section 3, we can assume that Bob can determine b with probability $1-2\varepsilon$ for x drawn uniformly from the set A, which is of size at least 2^{2n-c-1}. This means that he can distinguish whether his string w comes from a "noisy" Mx or from a "noisy" $Mx \oplus 1^n$. Recall that our hard distribution is uniform on the x's, the matchings M, and the bit b, and we pick w by independently choosing each bit w_ℓ such that $\Pr[w_\ell = (Mx)_\ell \oplus b] = 3/4$. Call $\mathcal{D}_{0M}$ the distribution on the strings w induced by our hard distribution when we condition on $b = 0$, on fixed matching M, and x is uniformly picked from A. Denote the corresponding distribution when $b = 1$ by $\mathcal{D}_{1M}$. The probability to distinguish two distributions of total variation distance d is at most $(1+d)/2$. Hence, since Bob has success probability at least $1-2\varepsilon$, the distributions $\mathcal{D}_{0M}$ and $\mathcal{D}_{1M}$ must be far apart on average:

$$\frac{1}{|\mathcal{M}|}\sum_{M\in\mathcal{M}} d(\mathcal{D}_{0M}, \mathcal{D}_{1M}) \geq 1 - 4\varepsilon, \tag{1}$$

where $\mathcal{M}$ is the set of all perfect matchings. Below, we upper bound the average $d(\mathcal{D}_{0M}, \mathcal{D}_{1M})$, which implies an upper bound on $|A|$ (and hence a lower bound on c).

To express $d(\mathcal{D}_{0M}, \mathcal{D}_{1M})$, we define the following probability distributions. Let μ be the distribution on a bit such that $\mu(0) = 3/4$ and $\mu(1) = 1/4$. For $b \in \{0,1\}$ define the product distributions on $\{0,1\}^n$ as $f_b(y) = \prod_{i=1}^n \mu(y_i \oplus b)$. In other words, f_0 is the distribution on n-bit strings where each bit is independently 0 with probability 3/4 and 1 with probability 1/4 and f_1 is the same distribution with bits flipped. They represent the "noise" added to z. Let

$$g_M(z) = \frac{|\{x \in A \mid Mx = z\}|}{|A|}.$$

The distribution $\mathcal{D}_{0M}$ can be viewed as first picking a string z according to g_M and then adding noise according to f_0. This can be expressed as the *convolution* of f_0 and g_M, i.e.

$$\Pr_{\mathcal{D}_{0M}}[w] = \sum_{z\in\{0,1\}^n} f_0(z \oplus w)\cdot g_M(z) = 2^n \cdot f_0 * g_M(w),$$

and similarly for $\mathcal{D}_{1M}$. This gives

$$d(\mathcal{D}_{0M}, \mathcal{D}_{1M}) = \frac{1}{2}\sum_{w\in\{0,1\}^n}\left|\Pr_{\mathcal{D}_{0M}}[w] - \Pr_{\mathcal{D}_{1M}}[w]\right| =$$

$$2^{n-1}\sum_{w\in\{0,1\}^n}|(f_0 - f_1) * g_M(w)| = 2^{2n}||\frac{f_0 - f_1}{2} * g_M||_1. \tag{2}$$

To get an upper bound on $d(\mathcal{D}_{0M}, \mathcal{D}_{1M})$, we upper bound the ℓ_1 norm by the ℓ_2 norm and use Parseval's identity (Lemma 4) to go to the Fourier domain:

$$||\frac{f_0 - f_1}{2} * g_M||_1^2 \leq$$
$$||\tfrac{f_0-f_1}{2} * g_M||_2^2 = \sum_{s\in\{0,1\}^n}\left(\widehat{\tfrac{f_0-f_1}{2}}(s)\right)^2 \cdot (\widehat{g_M}(s))^2 . \tag{3}$$

It is easy to see that the Fourier coefficients of $\frac{f_0-f_1}{2}$ are

$$\widehat{\frac{f_0 - f_1}{2}}(s) = \begin{cases} \frac{1}{2^{n+k}} & \text{for } s \text{ with } h(s) = k \text{ , } k \text{ odd} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Note that the parameter k denotes Hamming weight and takes integer values between 0 and $2n$. We now relate the uniform distribution on A to the Fourier coefficients of g_M, i.e. the distribution on the strings $z = Mx \in \{0,1\}^n$ induced by the matching M and by picking a uniform $x \in A$. Let $g : \{0,1\}^{2n} \to \mathbb{R}$ be the uniform distribution over the set A

$$g(x) = \begin{cases} \frac{1}{|A|} & \text{for } x \in A \\ 0 & \text{for } x \notin A \end{cases}$$

Note that for $x \in \{0,1\}^{2n}$ and $s \in \{0,1\}^n$ we have $(Mx)\cdot s = (xM^T)\cdot s = x\cdot(M^T s)$. By the definition of g_M,

$$\begin{aligned}
\widehat{g_M}(s) &= \frac{1}{2^n}\sum_{y\in\{0,1\}^n} g_M(y)(-1)^{y\cdot s} \\
&= \frac{1}{2^n|A|}\Big(|\{x \in A \mid (Mx)\cdot s = 0\}| - |\{x \in A \mid (Mx)\cdot s = 1\}|\Big) \\
&= \frac{1}{2^n|A|}\Big(|\{x \in A \mid x\cdot(M^T s) = 0\}| - |\{x \in A \mid x\cdot(M^T s) = 1\}|\Big) \\
&= \frac{1}{2^n}\sum_{x\in\{0,1\}^{2n}} g(x)(-1)^{x\cdot(M^T s)} = 2^n\cdot\hat{g}(M^T s).
\end{aligned} \tag{5}$$

Combining inequalities (1)- (5):

$$\begin{aligned}
(1-4\varepsilon)^2 \leq \frac{1}{|\mathcal{M}|}\sum_{M\in\mathcal{M}} d(\mathcal{D}_{0M}, \mathcal{D}_{1M})^2 &= \frac{2^{4n}}{|\mathcal{M}|}\sum_{M\in\mathcal{M}}||\frac{f_0 - f_1}{2} * g_M||_1^2 \\
&\leq \frac{2^{4n}}{|\mathcal{M}|}\sum_{M\in\mathcal{M}}\sum_{s\in\{0,1\}^n}\left(\widehat{\frac{f_0 - f_1}{2}}(s)\right)^2\cdot(\widehat{g_M}(s))^2 \\
&= \frac{2^{4n}}{|\mathcal{M}|}\sum_{M\in\mathcal{M}}\sum_{\substack{s:h(s)=k \\ k \text{ odd}}}\frac{1}{2^{2k}}\cdot\left(\hat{g}(M^T s)\right)^2.
\end{aligned} \tag{6}$$

Note that $h(M^T s) = 2h(s)$ and hence if $h(s)$ is odd, then $h(M^T s) = 2 \mod 4$. For $k = 2 \mod 4$ we define γ_k as follows: Let $v \in \{0,1\}^{2n}$ be a string of Hamming weight k and M be a random matching. Then $\gamma_k = \Pr_M[\exists s \text{ s.t. } v = M^T s]$. This probability depends only on k and we have

$$\begin{aligned}
&\sum_{\substack{s:h(s)=k \\ k \text{ odd}}}\frac{1}{2^{2k}}\frac{1}{|\mathcal{M}|}\sum_{M\in\mathcal{M}}\left(\hat{g}(M^T s)\right)^2 \\
&= \sum_{\substack{v:h(v)=2k \\ k \text{ odd}}}\frac{1}{2^{2k}}\,\gamma_{2k}\,(\hat{g}(v))^2 = \sum_{\substack{v:h(v)=k \\ k=2(mod4)}}\frac{1}{2^k}\,\gamma_k\,(\hat{g}(v))^2.
\end{aligned} \tag{7}$$

Call $\delta_k = \gamma_k^{1/k}$. Combining (6) and (7) we get

$$(1-4\varepsilon)^2 \le 2^{4n} \sum_{k=2(mod4)} \frac{1}{2^k} \sum_{v:h(v)=k} (\delta_k)^{h(v)} (\hat{g}(v))^2$$
$$\le 2^{4n} \sum_{k=2(mod4)} \frac{1}{2^k} \sum_{v\in\{0,1\}^{2n}} (\delta_k)^{h(v)} (\hat{g}(v))^2 \quad (8)$$

We can upper bound γ_k: for any even number $t \ge 2$, let $N(t)$ be the number of perfect matchings on $[t]$. Then, $N(2) = 1$, $N(t) = (t-1)N(t-2)$. It is not hard to see that the expression for γ_k is

$$\gamma_k = \frac{N(k)N(2n-k)}{N(2n)} \le \left(\frac{k}{2n}\right)^{k/2}.$$

Then $0 \le \delta_k \le \sqrt{\frac{k}{2n}} \le 1$ for $k \in [2, 2n]$. We now apply the KKL inequality (Lemma 5) to the function $g \cdot |A|$ (hence $t = |A|/2^{2n}$) and use $|A| \ge 2^{2n-c-1}$

$$\sum_{v\in\{0,1\}^{2n}} (\delta_k)^{h(v)} (\hat{g}(v))^2 \le \frac{1}{|A|^2} \left(\frac{|A|}{2^{2n}}\right)^{\frac{2}{1+\delta_k}}$$
$$\le 2^{-4n} \left(\frac{2^{2n}}{|A|}\right)^{2\delta_k} \le 2^{-4n+(c+1)\sqrt{\frac{2k}{n}}}.$$

Finally, combining with inequality (8) implies $(1-4\varepsilon)^2 \le$

$$\sum_{k=2(mod4)} 2^{-k+(c+1)\sqrt{\frac{2k}{n}}} = \sum_{k=2(mod4)} 2^{-k/2} \left(2^{-k/2+(c+1)\sqrt{\frac{2k}{n}}}\right).$$

Since $\sum_{k\in[0,2n],k=2(mod4)} 2^{-k/2} = \sum_{k\in[0,n],k\ odd} 2^{-k} \le \frac{2}{3}$ there is a k such that $\frac{3}{2}(1-4\varepsilon)^2 \le 2^{-k/2+(c+1)\sqrt{\frac{2k}{n}}}$. Hence $c \ge \frac{\sqrt{n}}{2} - 1$ for sufficiently small ε. □

4. LOWER BOUND FOR αPM

In our proof for αPM, we look again at the set A that contains at least 2^{2n-c-1} x's and has error at most 2ε. Now we prove a stronger statement. From Bob's point of view the following happens when he receives the message corresponding to the set A: a uniformly picked matching M of αn disjoint edges (i_ℓ, j_ℓ), $\ell \in [\alpha n]$, is given, and an unknown x is picked uniformly from A. As before, define $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$ and $z = z_1 \dots z_{\alpha n}$. Note that z is a function of x and M. Here Bob knows M and he knows that x is a uniformly chosen element from the known set A. Bob needs to figure out whether his second input w equals $z \oplus 0^{\alpha n}$ or $z \oplus 1^{\alpha n}$. We will use capital letters to denote the corresponding random variables. In Theorem 11, we show that Z is close to uniformly distributed when the edges are known but x is not: if the communication c is "small", then the total variation distance (conditioned on M) between Z and the uniform distribution $U_{\alpha n}$ on αn bits is $\mathrm{E}_M[d(Z, U_{\alpha n} \mid M)] =$

$$= \mathrm{E}_M \left[\frac{1}{2} \sum_{z\in\{0,1\}^{\alpha n}} \left|\Pr[Z = z \mid M] - 2^{-\alpha n}\right|\right] \le \eta$$

for some small η; the expectation is taken over uniform M. Then also $\mathrm{E}_M[d(Z \oplus 0^{\alpha n}, U_{\alpha n} \mid M)] \le \eta$ and $\mathrm{E}_M[d(Z \oplus 1^{\alpha n}, U_{\alpha n} \mid M)] \le \eta$, and hence $\mathrm{E}_M[d(Z \oplus 0^{\alpha n}, Z \oplus 1^{\alpha n} \mid M)] \le \mathrm{E}_M[d(Z \oplus 0^{\alpha n}, U_{\alpha n} \mid M)] + \mathrm{E}_M[d(Z \oplus 1^{\alpha n}, U_{\alpha n} \mid M)] \le 2\eta$. Distinguishing between the two distributions $Z \oplus 0^{\alpha n}$ and $Z \oplus 1^{\alpha n}$ is exactly what Bob needs to do to determine b. It is well known that distinguishing between two distributions with variation distance 2η can be done with probability at most $1/2 + \eta$. Accordingly, if c is "small" then the success probability will be close to $1/2$. Since Bob's success probability on the set A is at least $1 - 2\varepsilon$, c must be large.

In what follows we analyze the distribution of the αn-bit string Z and prove Theorem 11. The random variable Z depends on the known matching M with edges $e_1 = (i_1, j_1), \dots, e_{\alpha n} = (i_{\alpha n}, j_{\alpha n})$ as well as on the unknown x, which is uniformly drawn from set A. The typical case is where $|A| \approx 2^{2n-c}$. Intuitively, if c is small (i.e. A is large), then for most M and strings $z \in \{0,1\}^{\alpha n}$ we should have $\Pr[Z = z \mid M] \approx 2^{-\alpha n}$. Hence $d(Z, U_{\alpha n} \mid M)$ should be small for most M, and $\mathrm{E}_M[d(Z, U_{\alpha n} \mid M)]$ should be small as well. Proving this will be quite technical.

We view the edges of M as being picked one by one. Since A is quite large, for most (i,j)-pairs roughly equally many x's should have $x_i \oplus x_j = 1$ as have $x_i \oplus x_j = 0$. Thus we expect the first bit Z_1 to be close to uniformly distributed when x is picked uniformly from A. Similarly, we would like the later bits Z_ℓ to be more or less uniform when conditioned on values $Z_1 = z_1, \dots, Z_{\ell-1} = z_{\ell-1}$ for the earlier edges. More formally, once $(i_1, j_1), \dots, (i_{\ell-1}, j_{\ell-1})$ and $z_1, \dots, z_{\ell-1}$ have been fixed, we define the "ℓ-th bias" by

$$\beta_\ell = \Pr_{x\in A}[Z_\ell = 1 \mid Z_1 = z_1, \dots, Z_{\ell-1} = z_{\ell-1}, M] - 1/2.$$

This is a function of the first ℓ edges of M and of the first $\ell - 1$ bits of Z. Though we write 'M' in the conditional for brevity, β_ℓ is actually independent of the last $\alpha n - \ell$ edges of M. It is positive if Z_ℓ is biased towards 1, and negative if Z_ℓ is biased towards 0. Note that a fixed M, z pair fully determines all biases $\beta_1, \dots, \beta_{\alpha n}$ and $\Pr_{x\in A}[Z = z \mid M] =$

$$\prod_{\ell=1}^{\alpha n} \Pr_{x\in A}[Z_\ell = z_\ell | Z_1 = z_1, ., Z_{\ell-1} = z_{\ell-1}, M] = \prod_{\ell=1}^{\alpha n} \left(\frac{1}{2} - (-1)^{z_\ell}\beta_\ell\right)$$

Fixing the first $\ell - 1$ edges of M and conditioning on their bitvalues $Z_1 = z_1, \dots, Z_{\ell-1} = z_{\ell-1}$ will shrink the set of possible x's. Let A_ℓ be the subset of A that is still consistent. Initially we have $|A_1| = |A| \ge 2^{2n-c-1}$. When we pick the next edge (i_ℓ, j_ℓ) and its value z_ℓ, the new set $A_{\ell+1}$ will be smaller by a factor $1/2 + \beta_\ell$ if $z_\ell = 1$ and by a factor $1/2 - \beta_\ell$ if $z_\ell = 0$. We have $|A_\ell| =$

$$|A| \cdot \Pr_{x\in A}[Z_1 = z_1, .., Z_{\ell-1} = z_{\ell-1} | M] = |A| \cdot \prod_{i=1}^{\ell-1} \left(\frac{1}{2} - (-1)^{z_i}\beta_i\right)$$

Hence we expect the set to shrink by about two for each new edge and bitvalue for that edge ($|A_\ell| \ge 2^{2n-c-\ell}$).

We use a result of Talagrand [32] to relate the expected squared bias β_ℓ^2 to the size of the set A_ℓ. Talagrand himself derived this using a large deviation inequality from [20], but Oded Regev showed us how it can be obtained in a simple way from the KKL inequality.

LEMMA 6 ([32], EQ. (2.9)). *For every $A \subseteq \{0,1\}^{2n}$, with $\beta_{ij} = \Pr_{x\in A}[x_i \oplus x_j = 1] - 1/2$, we have*

$$\sum_{i,j\in[2n],i\ne j} \beta_{ij}^2 \le \left(\log\left(\frac{2^{2n}}{|A|}\right)\right)^2.$$

PROOF. Let $f : \{0,1\}^{2n} \to \{0,1\}$ be the characteristic function of our set A, and $t = |A|/2^{2n}$. Let $s_{ij} \in \{0,1\}^{2n}$

be the string having a 1 only at positions i and j. Then

$$\hat{f}(s_{ij}) = \frac{1}{2^{2n}} \sum_{y \in \{0,1\}^{2n}} f(y)(-1)^{y \cdot s_{ij}}$$

$$= \frac{|A|}{2^{2n}} \cdot \frac{|\{y \in A | y \cdot s_{ij} = 0\}| - |\{y \in A | y \cdot s_{ij} = 1\}|}{|A|} = 2t\beta_{ij}.$$

Applying KKL (Lemma 5) to f, for every $\delta \in [0,1]$ we have:

$$\sum_{i,j \in [2n], i \neq j} \delta^2 \hat{f}(s_{ij})^2 \leq \sum_{s \in \{0,1\}^{2n}} \delta^{h(s)} \hat{f}(s)^2 \leq t^{2/(1+\delta)}.$$

Hence

$$\sum_{i,j \in [2n], i \neq j} \beta_{ij}^2 = \frac{1}{4t^2} \cdot \frac{1}{\delta^2} \sum_{i,j \in [2n], i \neq j} \delta^2 \hat{f}(s_{ij})^2 \leq \frac{1}{4\delta^2} \cdot t^{-2\delta}.$$

Picking $\delta = 1/\log(1/t) = 1/\log(2^{2n}/|A|)$ gives the lemma. □

This will allow us to show that β_ℓ is probably quite small *if* the set A_ℓ hasn't shrunk too fast. We allow some more shrinking than we expect: note the '$3c$' instead of 'c' in the exponent below. The way to read this corollary is as follows: the first $\ell - 1$ edges of M and the first $\ell - 1$ bits of z have already been fixed. This determines the set A_ℓ, and we assume this set is large enough. Choosing the ℓ-th edge of M will now determine the value of the ℓ-th bias β_ℓ. The corollary bounds the expectation of β_ℓ^2, where the expectation is taken over all choices for the ℓ-th edge of M.

COROLLARY 7. *There is an absolute constant $\gamma > 0$ such that if $|A_\ell| \geq 2^{2n-3c-\ell}$, then*
(1) $E[\beta_\ell^2] \leq \gamma(c/n)^2$ and (2) $\Pr[|\beta_\ell| \geq \varepsilon] \leq \gamma(\frac{c}{n\varepsilon})^2$.

PROOF. Note that fixing a bitvalue for the parity of an edge means that the two bits in that edge behave as one bit. Accordingly, we can view the set A_ℓ as a set of strings of length $m = 2n-(\ell-1)$ bits. We can upper bound the sum of biases over all possible new edges (excluding ones touching earlier edges) by the sum over all possible edges (including ones touching earlier edges):

$$\sum_{i_\ell, j_\ell \in [m] \setminus \{i_1, \ldots, i_{\ell-1}, j_1, \ldots, j_{\ell-1}\}, i \neq j} \beta_{i_\ell, j_\ell}^2 \leq \sum_{i,j \in [m], i \neq j} \beta_{ij}^2 \leq O(c^2),$$

where the last inequality is by applying Lemma 6 to A_ℓ. Dividing by the number $\binom{2n-2(\ell-1)}{2} = \Theta(n^2)$ of possible new edges proves part (1). Part (2) now follows from Chebyshev's inequality. □

Note that on the one hand we need to assume that the sets A_ℓ are not too small in order to show that the biases β_ℓ are probably not too large (via Corollary 7). But on the other hand we need to show that the earlier biases are not too large in order to be able to conclude that A_ℓ is not too small. To deal with this problem, below we give a proof in two "passes". The first pass is quite coarse-grained and shows that (with high probability) the sets A_ℓ won't shrink by a factor of 2^{-2c} more than what we expect. Thus we will have $|A_\ell| \geq 2^{2n-3c-\ell}$ for each ℓ, which allows us to apply Corollary 7 to each of the αn biases during the second pass. In this second, more fine-grained pass we actually show that Z is close to uniformly distributed, conditioned on M.

4.1 First pass: A_ℓ probably don't shrink much

We can only use Corollary 7 if the condition $|A_\ell| \geq 2^{2n-3c-\ell}$ is satisfied. We now show that with high probability (over the uniform distribution on M, z) this is indeed the case for all ℓ simultaneously. The proof uses the following concentration result from [23].

LEMMA 8 ([23], THM. 3.7). *Let $S_1, \ldots, S_k$ be bounded random variables with $E[S_j | S_1 = s_1, \ldots, S_{j-1} = s_{j-1}] = 0$ for all $1 \leq j \leq k$ and all $s_1, \ldots, s_k$. Then for all $t, v \geq 0$*

$$\Pr\left[\sum_{j=1}^k S_j \geq t\right] \leq e^{-t^2/2v} + \Pr\left[\sum_{j=1}^k S_j^2 \geq v\right].$$

LEMMA 9. *Let $\eta \in [0,1]$ and $\alpha \leq \sqrt{\eta/256\gamma \log n}$. Suppose x is uniformly drawn from a set A of size at least 2^{2n-c-1}, where $c \leq \sqrt{\eta n/64\alpha\gamma}$. Then with probability at least $1-\eta$ (over uniformly chosen M, z) the following holds: for each $\ell \in [\alpha n]$ we have $|A_\ell| \geq 2^{2n-3c-\ell}$ and $|\beta_\ell| \leq 1/4$.*

PROOF. Assume $c = \sqrt{\eta n/64\alpha\gamma}$ for simplicity. Defining $S_i = -(-1)^{z_i} 2\beta_i$, we have

$$|A_\ell| == |A| \cdot \prod_{i=1}^{\ell-1} (1/2 - (-1)^{z_i}\beta_i) \geq 2^{2n-c-\ell} \prod_{i=1}^{\ell-1} (1 + S_i).$$

To lower bound $|A_\ell|$ it thus suffices to lower bound $\prod_{i=1}^{\ell-1}(1+S_i)$ by 2^{-2c} under distribution $\mathcal{P}$, which is uniform on the matching M and on z. Taking natural logarithms, we need to show for any ℓ

$$\ln\left(\prod_{i=1}^{\ell-1}(1+S_i)\right) = \sum_{i=1}^{\ell-1} \ln(1+S_i) \geq -c2\ln(2). \quad (9)$$

Let us divide the αn ℓ's into blocks of size c each: for $1 \leq k \leq \alpha n/c$ define the k-th block $B_k = \{(k-1)c+1, \ldots, kc\}$ (assume for simplicity that $\alpha n/c$ is an integer). Let E_k be the following event:

(a) $|\beta_i| \leq 1/4$ for each $i \in B_k$ and
(b) $\sum_{i \in B_k} \ln(1+S_i) \geq -c^2\ln(2)/\alpha n$.

We will show below in Claim 10 that for all $k \in [\alpha n/c]$, $\Pr_{\mathcal{P}}[\neg E_k \mid E_1, \ldots, E_{k-1}] \leq \frac{c}{\alpha n}\eta$. This implies

$$\Pr_{\mathcal{P}}[\neg(E_1, \ldots, E_{\alpha n/c})] \leq \sum_{k=1}^{\alpha n/c} \Pr_{\mathcal{P}}[\neg E_k \mid E_1, \ldots, E_{k-1}] \leq \eta.$$

If $E_1, \ldots, E_{\alpha n/c}$ all hold, then from (b) for all k we have

$$\sum_{i=1}^{k \cdot c} \ln(1+S_i) \geq -k \cdot c^2 \ln(2)/\alpha n \geq -c\ln(2)$$

and in particular Eq. (9) holds (even with righthand side of $-c\ln(2)$ instead of $-c2\ln(2)$) whenever $\ell - 1$ is a multiple of c. For the other ℓ, pick k such that $\ell - 1 \in B_{k+1}$ and note that thanks to (a) we have $\ln(1+S_i) \geq -\ln(2)$ and hence

$$\sum_{i=1}^{\ell-1} \ln(1+S_i) = \sum_{i=1}^{kc} \ln(1+S_i) + \sum_{i=kc+1}^{\ell-1} \ln(1+S_i)$$

$$\geq -c\ln(2) + \sum_{i=kc+1}^{\ell-1} -\ln(2) \geq -c2\ln(2).$$

It thus remains to prove

CLAIM 10. $\Pr_{\mathcal{P}}[\neg E_k \mid E_1, \ldots, E_{k-1}] \le \frac{c}{\alpha n}\eta$ *for all* k.

PROOF. We have $\Pr_{\mathcal{P}}[\neg E_k \mid E_1, \ldots, E_{k-1}] \le \Pr_{\mathcal{P}}[\neg(a) \mid E_1, \ldots, E_{k-1}] + \Pr_{\mathcal{P}}[\neg(b) \mid E_1, \ldots, E_{k-1}, (a)]$. We bound the two terms on the righthand side separately, starting with the first.

Let $\ell_1 = (k-1)c$ be the last index in B_{k-1}. Conditioning on $E_1, \ldots, E_{k-1}$ means that $|A_{\ell_1+1}| \ge 2^{2n-2c-\ell_1-1}$. For each $i \in B_k$, if $|\beta_{\ell_1+1}|, \ldots, |\beta_{i-1}| \le 1/4$ then as before $\sum_{j=\ell_1+1}^{i-1} \ln(1+S_j) \ge -c\ln(2)$ and hence $|A_i| \ge |A_{\ell_1+1}| \cdot 2^{-(i-\ell_1-1)-c} \ge 2^{2n-3c-i}$. By Corollary 7 (part 2) we have

$$\Pr_{\mathcal{P}}[|\beta_i| > 1/4] \le \gamma\left(\frac{4c}{n}\right)^2 = \frac{\eta}{4\alpha n}$$

and hence (a) fails to hold for block k with probability

$$\Pr_{\mathcal{P}}[\neg(a) \mid E_1, \ldots, E_{k-1}] \le \frac{\eta c}{4\alpha n}$$

Next, conditioning on (a) and $E_1, \ldots, E_{k-1}$ we show that (b) holds for B_k with probability at least $1 - 3\eta c/4\alpha n$, which implies the claim. Let $\mathcal{P}'$ be the distribution on the edges and the string z when we condition on (a). We make some observations about $\mathcal{P}'$. First, like $\mathcal{P}$, we can view $\mathcal{P}'$ as picking edges and bits z_i sequentially: select the i-th edge uniformly at random among all edges that are disjoint from those already chosen *and* that have bias $\le 1/4$; then pick z_i uniformly at random. The difference with $\mathcal{P}$ is that the i-th edge is not picked arbitrarily, but is restricted to edges having bias $\le 1/4$. Second, the condition of Lemma 8 holds for each $S_i = -(-1)^{z_i}2\beta_i$: the conditional expectations are all 0, because we first determine β_i and then give S_i the sign $+$ or $-$ with equal probability.

Since $|S_i| = 2|\beta_i| \le 1/2$ for $i \in B_k$, we have $\ln(1+S_i) \ge S_i - S_i^2$, and hence

$$\sum_{i\in B_k} \ln(1+S_i) \ge \sum_{i\in B_k} S_i - \sum_{i\in B_k} S_i^2.$$

Let $v = c^2\ln(2)/2\alpha n$; this is half of what (b) allows us to lose (note that $v \ge 2\ln n$ by our choice of parameters). Then,

$$\begin{aligned}
\Pr_{\mathcal{P}'}[\neg(b) \mid E_1, \ldots, E_{k-1}] &= \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} \ln(1+S_i) < -2v\right] \\
&\le \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i - \sum_{i\in B_k} S_i^2 < -2v\right] \\
&\le \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i < -v\right] + \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i^2 > v\right]. \qquad (10)
\end{aligned}$$

First we bound the second term of the righthand side of Eq. (10). Corollary 7 implies, both under $\mathcal{P}$ and $\mathcal{P}'$:

$$\mathrm{E}\left[\sum_{i\in B_k} S_i^2\right] = \sum_{i\in B_k} 4\mathrm{E}\left[\beta_i^2\right] \le c\cdot 4\gamma\left(\frac{c}{n}\right)^2 = \frac{4\gamma c^3}{n^2},$$

By Markov's inequality

$$\Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i^2 > v\right] \le \frac{4\gamma c^3}{vn^2} = \frac{8\alpha\gamma}{\ln 2}\frac{c}{n} \le \frac{\eta c}{4\alpha n},$$

where the equality follows from our value of c, and the last inequality follows easily from our upper bound on α.

Now we bound the first term on the right of Eq. (10). By Lemma 8 (with $t = v \ge 2\ln n$),

$$\Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i < -v\right] \le e^{-v/2} + \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i^2 \ge v\right] \le \frac{\eta c}{2\alpha n}.$$

Putting everything together:

$$\begin{aligned}
&\Pr_{\mathcal{P}}[\neg E_k \mid E_1, \ldots, E_{k-1}] \\
&\le \Pr_{\mathcal{P}}[\neg(a) \mid E_1, \ldots, E_{k-1}] + \Pr_{\mathcal{P}}[\neg(b) \mid E_1, \ldots, E_{k-1}, (a)] \\
&= \Pr_{\mathcal{P}}[\neg(a) \mid E_1, \ldots, E_{k-1}] + \Pr_{\mathcal{P}'}[\neg(b) \mid E_1, \ldots, E_{k-1}] \\
&\le \frac{\eta c}{4\alpha n} + \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i < -v\right] + \Pr_{\mathcal{P}'}\left[\sum_{i\in B_k} S_i^2 > v\right] \\
&\le \frac{\eta c}{4\alpha n} + \frac{\eta c}{2\alpha n} + \frac{\eta c}{4\alpha n} = \frac{\eta c}{\alpha n}.
\end{aligned}$$

This concludes the proof of Claim 10. □

This concludes the proof of Lemma 9.

4.2 Second pass: Z is close to uniform

We now prove the main result about αPM.

THEOREM 11. *Let* $\eta \in [0,1]$ *and* $\alpha \le \sqrt{\eta/256\gamma\log n}$. *Suppose* x *is uniformly drawn from a set* A *of size at least* 2^{2n-c-1}, *where* $\le \sqrt{\eta^3 n/2^{14}\ln(64/\eta)\alpha\gamma} = O(\sqrt{\eta^3 n/\alpha})$, *then* $E_M[d(Z, U_{\alpha n} \mid M)] \le \eta$.

PROOF. Let $S_\ell = -(-1)^{z_\ell}2\beta_\ell$ and $v = \eta^2/32\ln(64/\eta)$. Call a pair M, z "good" if the following three things hold for it: (1) $|S_\ell| \le 1/2$ for all ℓ, (2) $\sum_{\ell=1}^{\alpha n} S_\ell^2 \le v$, and (3) $|\sum_{\ell=1}^{\alpha n} S_\ell| \le \eta/4$. Call the pair M, z "bad" otherwise.

Letting $\#M$ be the number of α-matchings M, we rewrite the expected total variation distance as:

$$\begin{aligned}
\mathrm{E}_M\left[d(Z, U_{\alpha n}|M)\right] &= \frac{1}{2\#M}\sum_{M,z}\left|\Pr_{x\in A}[Z = z \mid M] - 2^{-\alpha n}\right| \\
&= \frac{1}{2\#M}\sum_{\mathrm{good}\,M,z}\left|\Pr_{x\in A}[Z = z \mid M] - 2^{-\alpha n}\right| \\
&\quad + \frac{1}{2\#M}\sum_{\mathrm{bad}\,M,z}\left|\Pr_{x\in A}[Z = z \mid M] - 2^{-\alpha n}\right|.
\end{aligned}$$

Let $\mathcal{P}$ be the uniform distribution on M, z. We start by bounding the probability (over $\mathcal{P}$) that M, z is a bad pair. Since $c \le \sqrt{\eta^3 n/2^{14}\ln(64/\eta)\alpha\gamma} \le \sqrt{(\eta/128)n/64\alpha\gamma}$, we can apply Lemma 9 with value $\eta/128$ for η. Let B denote the bad event that at least one A_ℓ is too small and C the bad event that at least one S_ℓ has absolute value larger than $1/2$. Then by Lemma 9 $\Pr_{\mathcal{P}}[B] \le \eta/128$ and $\Pr_{\mathcal{P}}[C] \le \eta/128$. From Corollary 7 we have

$$\mathrm{E}_{\mathcal{P}}\left[\sum_{\ell=1}^{\alpha n} S_\ell^2 \,\middle|\, \neg B\right] \le 4\alpha n\gamma(c/n)^2 \le \eta v/128.$$

By Markov, $\Pr_{\mathcal{P}}\left[\sum_{\ell=1}^{\alpha n} S_\ell^2 \geq v \middle| \neg B\right] \leq \eta/128$, hence

$$\Pr_{\mathcal{P}}\left[\sum_{\ell=1}^{\alpha n} S_\ell^2 \geq v\right] \leq \Pr_{\mathcal{P}}[B] + \Pr_{\mathcal{P}}\left[\sum_{\ell=1}^{\alpha n} S_\ell^2 \geq v \middle| \neg B\right] \leq \eta/64.$$

We now apply Lemma 8 with $t = \eta/4$ and our $v = \eta^2/32\ln(64/\eta)$ to show $\sum_{\ell=1}^{\alpha n} S_\ell$ is usually small (note $t^2/2v = \ln(64/\eta)$):

$$\begin{aligned}\Pr_{\mathcal{P}}\left[\left|\sum_{\ell=1}^{\alpha n} S_\ell\right| > \eta/4\right] &\leq 2\left(e^{-t^2/2v} + \Pr_{\mathcal{P}}\left[\sum_{\ell=1}^{\alpha n} S_\ell^2 \geq v\right]\right) \\ &\leq 2(\eta/64 + \eta/128) = 6\eta/128.\end{aligned}$$

Applying the union bound, we see that the probability that M, z is a bad pair is at most $(\eta/128+\eta/64+6\eta/128) < \eta/12$. For any pair M, z we have $\left|\Pr_{x\in A}[Z = z \mid M] - 2^{-\alpha n}\right| =$

$$\left|\prod_{\ell=1}^{\alpha n}(1/2 - (-1)^{z_\ell}\beta_\ell) - 2^{-\alpha n}\right| = 2^{-\alpha n}\left|\prod_{\ell=1}^{\alpha n}(1+S_\ell) - 1\right|.$$

Next, we show that $\left|\prod_{\ell=1}^{\alpha n}(1+S_\ell) - 1\right| \leq \eta/2$ for good pairs M, z. First, $\prod_{\ell=1}^{\alpha n}(1+S_\ell) \leq e^{\sum_{\ell=1}^{\alpha n} S_\ell} \leq e^{\eta/4} \leq 1 + \eta/2$. Second, since $|S_\ell| \leq 1/2$ for all ℓ, we have

$$\sum_{\ell=1}^{\alpha n}\ln(1+S_\ell) \geq \sum_{\ell=1}^{\alpha n} S_\ell - \sum_{\ell=1}^{\alpha n} S_\ell^2 \geq -\eta/4 - v$$

and $\prod_{\ell=1}^{\alpha n}(1+S_\ell) \geq e^{-\eta/4-v} \geq 1 - \eta/4 - v \geq 1 - \eta/2$. Hence, for good pairs M, z

$$\frac{1}{\#M}\sum_{\text{good}M,z}\left|\Pr_{x\in A}[Z = z|M] - 2^{-\alpha n}\right| \leq \frac{1}{\#M}\sum_{\text{good}M,z}\frac{\eta}{2}\cdot 2^{-\alpha n} \leq \frac{\eta}{2}.$$

Moreover, using also that the probability that M, z is a good pair is at least $1 - \eta/12$, we have

$$\frac{1}{\#M}\sum_{\text{good}M,z}\Pr_{x\in A}[Z = z \mid M] \geq (1-\eta/12)(1-\eta/2) \geq 1-7\eta/12$$

and therefore $\frac{1}{\#M}\sum_{\text{bad}M,z}\left|\Pr_{x\in A}[Z = z \mid M] - 2^{-\alpha n}\right|$

$$\leq \frac{1}{\#M}\sum_{\text{bad}M,z}\left(\Pr_{x\in A}[Z = z \mid M] + 2^{-\alpha n}\right) \leq \frac{7\eta}{12} + \frac{\eta}{12} = \frac{2\eta}{3}.$$

Now we can finally bound the expected total variation distance over all matchings: $\mathrm{E}_M\left[d(Z, U_{\alpha n} \mid M)\right] =$

$$\frac{1}{2\#M}\sum_{M,z}\left|\Pr_{x\in A}[Z = z \mid M] - 2^{-\alpha n}\right| \leq \frac{1}{2}\left(\frac{\eta}{2} + \frac{2\eta}{3}\right) \leq \eta. \quad \square$$

Note that the above theorem works for $\alpha = O(\sqrt{\eta/\log n})$. We would like to make it work also for constant α. Recently, we learned of a simplified proof of the lower bound on αPM by Oded Regev, who used the Fourier methods from our proof of the lower bound on NPM together with the Bonami-Beckner inequality, which may resolve this.

4.3 Consequences

As we explained in Section 3, a one-way protocol with c bits of communication and error probability ε implies the existence of a set $A \subseteq \{0,1\}^{2n}$ of size at least 2^{2n-c-1}, such that the protocol's error for a uniformly chosen $x \in A$ and matching M is at most 2ε. But then $\mathrm{E}_M[d(Z, U_{\alpha n} \mid M)]$ must have been large. Hence applying Theorem 11 with small constant η gives the lower bound on the classical communication c required to compute αPM. Combining that classical lower bound with the bounds mentioned in Section 2, we obtain the separation stated in the introduction:

Theorem 2. *For $\alpha \in [0, O(1/\sqrt{\log n})]$, the classical bounded-error one-way communication complexity of α-Partial Matching is $R^1_\varepsilon(\alpha\mathrm{PM}) = \Theta(\sqrt{n/\alpha})$; the quantum bounded-error one-way complexity is $Q^1_\varepsilon(\alpha\mathrm{PM}) = O(\log(n)/\alpha)$.*

The link with the bounded-storage model should be clear. Alice's input plays the role of the uniformly distributed bit-string X that is temporarily publicly available. The message plays the role of the adversary's memory, where he stores c classical bits that depend on X. Bob's matching plays the role of the initial shared secret key Y. Knowing Y, the two parties can compute the αn-bit string $Z(X, Y)$ from X, even in an online fashion. The fact that from Bob's perspective the string Z is close to uniform if Alice's message is too short, corresponds to the fact that the memory-bounded adversary in the cryptographic protocol knows hardly anything about $Z(X, Y)$—even if the adversary learns Y after X ceases to be public.

More precisely, assume w.l.o.g that the classical adversary is deterministic and stores some c-bit function $m(X)$ of uniformly distributed $2n$-bit string X, where $c = \sigma\sqrt{\eta^3 n/\alpha}$ (for sufficiently small σ). As before, this partitions the set of all x into 2^c sets. Almost all x sit in "large" sets: at most an $\eta/2$-fraction of the x sit in sets of size at most $2^{2n-c-\log(2/\eta)}$. Hence whenever the adversary has stored a c-bit string corresponding to a large set (which happens with probability $1 - \eta/2$), we can apply Theorem 11 to guarantee that the αn-bit string $Z(X, Y)$ is $\eta/2$-close to uniform, where Y is the uniformly chosen matching. Overall, $Z(X, Y)$ is η-close to uniform (even if Y is leaked afterwards).

In contrast, the quantum upper bound from Section 2 shows that storing $O(\log(n)/\alpha)$ qubits about X gives the adversary one bit of the string $Z(X, Y)$, when later Y is revealed. Hence the distance between Z and the uniform distribution is at least $1/2$. Using k times as much memory gives the adversary roughly k bits of $Z(X, Y)$, and the resulting distribution will be $1 - 2^{-k}$ far from uniform. This gives the theorem stated in the introduction:

Theorem 3. *Let $\eta \in [0, 1]$ and $\alpha \in [0, O(\sqrt{\eta/\log n})]$. The extracted αn-bit string in the bounded-storage protocol derived from the α-Partial Matching problem is η-secure against a classical adversary with memory bound $O(\sqrt{\eta^3 n/\alpha})$, while for every positive integer $k \ll \alpha n$ it is at most $(1 - 2^{-k})$-secure against an adversary with $O(k\log(n)/\alpha)$ qubits.*

So far, we have proved that if the $2n$-bit string X is uniformly distributed over a set A with $|A| \geq 2^{2n-c-1}$ (i.e., a flat distribution on A), then $Z(X, Y)$ is close to uniform even knowing Y, where Y is the uniformly chosen matching. By a result of Chor and Goldreich [10, Lemma 5] based on the fact that any distribution can be thought of as a convex combination of flat distributions, we can conclude that the same result is true in the more general situation when X has min-entropy greater than $2n - c - 1$ and hence conclude the result about extractors mentioned in Section 1.2.2.

Similarly, in the setting of privacy amplification, Alice's input plays the role of the shared random variable X, the

message in the communication protocol plays the role of the prior partial information that the adversary has about X (which is upper bounded by the communication c) and Bob's matching plays the role of the independent uniform seed Y. The fact that Bob's view about the string Z (which is the new secret key) is close to uniform when Alice's message is short, means that any adversary that has full knowledge of Y and partial information about X bounded by c, knows hardly anything about the string $Z(X,Y)$.

Acknowledgments. We thank Oded Regev, Guy Kindler, Jaikumar Radhakrishnan, Shengyu Zhang, Renato Renner, Yevgeniy Dodis, Christian Schaffner, Barbara Terhal and Scott Aaronson for many helpful discussions.

5. REFERENCES

[1] S. Aaronson. The learnability of quantum states. quant-ph/0608142, 18 Aug 2006.

[2] Y. Aumann, Y. Z. Ding, and M. Rabin. Everlasting security in the bounded storage model. *IEEE Trans. of Information Theory*, 48:1668–1680, 2002.

[3] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and Systems Sciences*, 58(1):137–147, 1999.

[4] C. H. Bennett, G. Brassard, and J.-M. Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, 1988.

[5] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *30th ACM STOC*, pages 63–68, 1998.

[6] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16), September 26, 2001.

[7] W. Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*,102:159–182,1975.

[8] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. In *36th ACM STOC*, pages 128–137, 2004.

[9] A. Bonami. Etude des coefficients de Fourier des fonctions de Lp(G). *Annales de l'Institute Fourier*, 20(2):335–402, 1970.

[10] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[11] S. Dziembowski and U. Maurer. Optimal randomizer efficiency in the bounded-storage model. *Journal of Cryptology*, 17(1):5–26, 2004.

[12] F. Le Gall. Exponential separation of quantum and classical online space complexity. In *18th ACM SPAA*, pages 67–73, 2006.

[13] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *38th ACM STOC*, pages 594–603, 2006.

[14] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973. English translation in *Problems of Information Transmission*, 9:177–183, 1973.

[15] R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *21st ACM STOC*, pages 12–24, 1989.

[16] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *29th IEEE FOCS*, pages 68–80, 1988.

[17] H. Klauck. Lower bounds for quantum communication complexity. In *42nd IEEE FOCS*, pages 288–297, 2001.

[18] E. Kushilevitz, N. Nisan. *Communication Complexity*. Cambridge Univ. Press, 1997.

[19] R. König and B. M. Terhal. The bounded storage model in the presence of a quantum adversary, 11 Aug 2006. quant-ph/0608101.

[20] M. Ledoux and M. Talagrand. *Probability in Banach Spaces*. Springer, 1991.

[21] C.-J. Lu. Encryption against storage-bounded adversaries from on-line strong extractors. *Journal of Cryptology*, 17(1):27–42, 2004.

[22] U. Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5(1):53–66, 1992.

[23] C. McDiarmid. Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, pages 195–248. Springer, Berlin, 1998.

[24] M. Muthukrishnan. *Data Streams:Algorithms and Applications*. Now Publishers, 2005.

[25] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[26] I. Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991.

[27] R. Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5(3/4):205–221, 1995.

[28] R. Raz. Exponential separation of quantum and classical communication complexity. In *31st ACM STOC*, pages 358–367, 1999.

[29] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zürich, 2005.

[30] R. Renner and R. König. Universally composable privacy amplification against quantum adversaries. In *2nd TCC*, vol. 3378 of *LNCS*, pages 407–425, 2005.

[31] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[32] M. Talagrand. How much are increasing sets correlated? *Combinatorica*, 16(2):243–258, 1996.

[33] S. P. Vadhan. Constructing locally computable extractors and cryptosystems in the bounded-storage model. *Journal of Cryptology*, 17(1):43–77, 2004.

[34] R. de Wolf. Quantum communication and complexity. *Theoretical Computer Science*, 287(1):337–353, 2002.

[35] A. C-C. Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th IEEE FOCS*, pages 222–227, 1977.

[36] A. C-C. Yao. Some complexity questions related to distributive computing. In *11th ACM STOC*, pages 209–213, 1979.